





Method for identifying and isolating genome fragments with coupling disequilibrium

Patent number: DE10017675
Publication date: 2001-12-06
Inventor: OLEK KLAUS (DE); HANSEN THOMAS (DE); WEBER JUERGEN (DE); LEUER MARCO (DE)
Applicant: QTL AG GES ZUR ERFORSCHUNG KOM (DE)
Classification:
- **international:** C12N15/10; C12N1/21
- **european:** C12N15/10; C12N15/10C
Application number: DE20001017675 20000408
Priority number(s): DE20001017675 20000408

Also published as:

 WO0177374 (A3)
 WO0177374 (A2)
 US2003165926 (A1)
 CA2416118 (A1)

[Report a data error here](#)

Abstract not available for DE10017675

Abstract of corresponding document: **US2003165926**

The subject of the invention is a method for the identification and isolation of genome fragments with coupling disequilibrium, wherein regions of the genome are isolated, which contain candidate gene regions that are found in coupling disequilibrium with their constricted DNA environment, and these genome regions are obtained from individuals who are not related to one another as well as individuals who are related to one another. The earlier conducting of the cloning step, in comparison to other methods, makes possible an obtaining of DNA fragments that is independent of quantity and in addition replaces the methylation step conducted in other methods. It was also found that the use of a plant enzyme according to the invention brings about a much higher specificity of the method when compared with those in which the Mut S, H, L complex is used.

Data supplied from the **esp@cenet** database - Worldwide

⑮ **BUNDESREPUBLIK
DEUTSCHLAND**



**DEUTSCHES
PATENT- UND
MARKENAMT**

⑫ **Offenlegungsschrift**
⑩ **DE 100 17 675 A 1**

⑤① Int. Cl.⁷:
C 12 N 15/10
C 12 N 1/21

⑳ Aktenzeichen: 100 17 675.5
㉔ Anmeldetag: 8. 4. 2000
㉓ Offenlegungstag: 6. 12. 2001

㉑ **Anmelder:**
QTL AG Gesellschaft zur Erforschung komplexer
genetischer Merkmale, 53359 Rheinbach, DE

㉒ **Vertreter:**
Schubert, K., Dipl.-Chem. Dr.rer.nat., Pat.-Anw.,
10119 Berlin

㉕ **Erfinder:**
Olek, Klaus, Prof. Dr., 53359 Rheinbach, DE;
Hansen, Thomas, Dipl.-Chem., 41065
Mönchengladbach, DE; Weber, Jürgen, Dr., 56567
Neuwied, DE; Leuer, Marco, Dipl.-Biol., 53111 Bonn,
DE

Die folgenden Angaben sind den vom Anmelder eingereichten Unterlagen entnommen

Prüfungsantrag gem. § 44 PatG ist gestellt

⑤④ **Verfahren zur Identifizierung und Isolierung von Genomfragmenten mit Kopplungsungleichgewicht**

⑤⑦ Gene, die polygen vererbte Merkmale steuern, sind im Gegensatz zu den Verursachergenen von monogen vererbten Merkmalen nicht wirklich effektiv mit Hilfe der Kopplungsstrategie zu identifizieren und zu isolieren. Dieses Verfahren ist zu teuer, zu ungenau und man benötigt zu viele Testindividuen. Das 1993 zum ersten Male beschriebene Prinzip des Genomic Mismatch Scanning (GMS) wird als eine Alternative angesehen, die all diese Nachteile nicht aufweist. Teil dieses GMS-Protokolls ist ein Selektionsschritt, der vom Prinzip her nur die Isolierung von Sequenzen gestattet, die ein bestimmtes Muster tragen (Methylierung). Ein anderer enzymatischer Schritt der Prozedur sehr unspezifisch (Enzym-Komplex Mut S, H, L). Diese und einige andere Zwischenschritte führen zu einer extrem niedrigen Ausbeute, die das nötige Weiterverarbeiten der Isolationsprodukte vor allem in Zusammenhang mit den schwierigen polygenen Merkmalen ungeheuer erschwert. Durch einen frühen Klonierungsschritt machen wir uns bei der Gewinnung das DNA-Fragment mengenunabhängig und ersetzen die Methylierung. Ersatz des Mut S, H, L-Komplexes durch ein Pflanzenenzym bewirkt eine weitaus höhere Spezifität des Verfahrens.

DE 100 17 675 A 1

DE 100 17 675 A 1

Beschreibung

Gebiet der Erfindung

- 5 [0001] Moderne Tierzucht macht sich seit einigen Jahren die molekularbiologische Analytik zu Nutze. Große Hoffnungen knüpfte man an die Strategie der markerunterstützten Züchtung. Zielgebiet dieser Bemühungen sind neben den klinisch und genetisch oft genau beschriebenen monogen vererbten Krankheiten die wirtschaftlich hochattraktiven QTL's. Es handelt sich dabei um sogenannte multifaktorielle Merkmale, das bedeutet, sie werden gleichzeitig durch mehrere genetische, aber eben auch äußere Einflußfaktoren gesteuert. Es sind im weitesten Sinne Leistungsparameter der Tierzucht.
- 10 Die vordergründig attraktivsten sind etwa Milchleistung, Fleischmenge, Wachstum und Fertilität. In diesen ersten 5 Jahren weltweiter systematischer Anwendung der markerunterstützten Züchtung sind die Ergebnisse eher enttäuschend; die Methode hat offensichtlich enge prinzipielle Grenzen, abgesehen von der Tatsache, daß sie ungeheuer aufwendig ist. Hier wird ein alternatives Verfahren vorgestellt.

Stand der Technik

- 15 [0002] Den seit langem ausgeübten Selektionsstrategien der Nutztierzucht steht seit einiger Jahren eine auf molekularbiologischen Methoden beruhende Alternative gegenüber, die man Markerunterstützte Selektion nennt (im englischen Sprachgebrauch: marker-assisted-selection oder abgekürzt MAS). Es handelt sich dabei im Kern um die Anwendung der
- 20 Kopplungsstrategie mit polymorphen DNA-Markern. Letztere hatte sich bei der Erforschung monogener Erbkrankheiten des Menschen als sehr erfolgreich erwiesen. Die Strategie war systematisch weltweit als eine erste Stufe des Human Genome Programms (HGP) angewendet worden. Seine theoretische Grundlage waren die Betrachtungen von Botstein et al 1980. Die enormen Anstrengungen des HGP lieferten schließlich die technische und intellektuelle Basis dafür, daß für die wichtigen Nutztierspezies Rind, Schwein und Schaf eine Markerkarte erstellt werden konnte. Im Rindengenom stehen nun rund 2000 lokalisierte polymorphe Marker zur Verfügung. Im Schweinegenom sind es mehr als 1200. Es handelt sich dabei meistens um Mikrosatelliten. Mit einem Sortiment von 300 solcher Marker, die möglichst gleichmäßig über das ganze Genom verteilt sind, hat man ein Instrument zur Verfügung, welches praktisch jedes monogen vererbte und phänotypisch klar beschriebene Merkmal zu lokalisieren gestattet. Dafür muß man mit diesem Markersatz mehrere hundert betroffene Tiere typisieren, das heißt man muß für die Tiere alle 300 Markergenloci analysieren. Mit Hilfe einer
- 25 ziemlich aufwendigen Mathematik findet man in aller Regel den Ort des Verursachergens im Genom. Ein typisches Ergebnis ist: Man hat den Genort auf 20–25 cM das entspricht 20–25 Megahasen eingeeengt. Als ein zum Teil schon praktisch verwertbares Ergebnis findet man zudem, daß innerhalb einer Familie bestimmte Allele der am nächsten gelegenen Markerloci zusammen mit dem interessanten Phänotyp vererbt werden. In einer anderen Familie kann es ein anderes Allel des gleichen Markergenortes sein. Innerhalb einer Familie kann man diesen Umstand schon diagnostisch nutzen, denn
- 30 in dieser einen Familie gestattet ein solches Allel genetische Vorhersagen über den zu erwartenden Phänotyp. Eine solche Aussage ist aber mit einer erheblichen Unsicherheit behaftet, diese Unsicherheit wird verursacht durch das Phänomen der Rekombination zwischen zwei Genorten. Die Unsicherheit kann man verringern, indem man Marker sucht, die näher beim Verursachergen liegen. In aller Regel wird man auf dieser Suche schließlich bei Markergenorten landen, die unmittelbar am oder im Gen selbst liegen. Zwischen diesen Markergenorten und dem Gen passieren praktisch keine Rekombinationen mehr. Läßt sich der interessierende Gendefekt auf einen Vorfahren zurückverfolgen, ist also dieser Defekt bei einem einzigen Vorfahren passiert und hat dieser sich von dort aus auf eine bestimmte Population verbreitet, so wird man die am und im Gen beobachtete Markerallelkombination in identischer Form bei allen betroffenen Genen dieser Population finden. Man sagt, die Marker befinden sich im Kopplungsungleichgewicht mit dem Verursachergen. Eine solche Markerallelkombination ist von außerordentlicher praktischer Bedeutung, denn über Familiengrenzen hinweg gestattet sie die erwünschte Voraussage des Phänotyps. Leider ist dieser Marsch von der 25 mB großen Kopplungsgruppe bis hin zum Gen ein sehr mühseliger; im Grenzfall dauert er Jahre.

- 35 [0003] Die Strategie ist auch auf polygen vererbte Merkmale anwendbar z. B. auf die sogenannten QTL's der Nutztierzucht. Die Situation ist nun aber wesentlich schwieriger, nicht nur weil man es mit mehreren Verursachergenen zu tun hat, bei denen man den Weg zum Gen bzw. Kopplungsungleichgewicht gehen muß – man erinnere sich, 25 cM das bedeutet 25 Millionen Basenpaare müssen studiert werden –, sondern weil die einzelnen Gene ganz unterschiedlich starken Einfluß auf den Phänotyp haben und weil dem auf die Genetik entfallende Einfluß auf den Phänotyp ein starker Umwelteinfluß gegenübersteht. So ist das bisherige Ergebnis einiger notgedrungen sehr groß angelegter QTL-Studien mit jeweils oft mehreren Tausend Tieren die Identifikation mehrerer Kopplungsgruppen, die meist 30–40 cM groß sind. Diagnostischen, aber damit auch züchterischen Wert haben sie im Grunde nur in Familien. Das bedeutet eine enorme Einschränkung der MAS. Das letztendliche Ziel all dieser molekularbiologischen Anstrengungen, nämlich die Isolierung der für die QTL's verantwortlichen Gene ist von dieser Basis aus überhaupt praktisch nicht zu erreichen.

- 45 [0004] Nachdem experimentell in zahlreichen Arbeitsgruppen belegt werden konnte, daß man mit der Kopplungsstrategie nicht nur analog zur Arbeit im menschlichen Genom monogen vererbte Krankheitsgene lokalisieren und dadurch züchterisch verwertbar machen konnte sondern auch die genetisch komplexen QTL's im Genom lokalisieren konnte, wurden die großen internationalen Züchtungsprogramme auf die zu erwartenden Ergebnisse der molekularen QTL-Suche eingestellt; wobei fast ausschließlich die gegenüber der konventionellen Strategie extrem verkürzte Zeitspanne bis zum Einsatz des als getestet anzusehenden Bullen in den Mittelpunkt der züchterischen Betrachtung gestellt wurde (Georges M 1998).

- 50 [0005] Die großen QTL-Suchprogramme – und wirklich kleine gibt es nicht wegen der hohen Anforderungen an die Zahl der zu untersuchenden Tiere – laufen fast alle nach einem Schema ab. Man nennt es Granddaughterdesign (GDD) (Geldermann 1975, Weller et al 1990).

- 55 [0006] Hierbei werden nur die beiden ersten Generationen genotypisiert, während die Phänotypisierung in der dritten Generation stattfindet. Eine typische Zahl für die zu untersuchenden Tiere ist: 10–20 Bullen, 50–100 Söhne von jedem

dieser Bullen, 50–100 Töchter pro Sohn. Bis zu 250.000 Genotypanalysen sind dabei erforderlich. Daß dieses Prinzip funktioniert, ist zum ersten Male von M. Georges gezeigt worden (Georges et al 1995). Die Studie wurde bei Genmark (USA) mit 1.518 Bullen, in 14 Holstein-Familien mit 159 Markern durchgeführt.

[0007] In Tabelle 1 sind die 1998 in Arbeit befindlichen GDD-Studien aufgeführt.

Tab. 1

Land	Organisation	Rasse	Familien	Söhne	Marker	Merkmal
USA	Genmark	Holstein	14	1518	159	P
USA	ABS	Holstein	14	1794	224	P, SCC, L
USA	Illinois Univ.	Holstein	18	1555	100	P
Israel	Israel Cattle					
USA	DBDR, USDA	Holstein	7	900	16	P, SCC
Kanada	Univ. Guelph	Holstein	7	450		P
NL	Holl. Genetics	Holstein	20	715	200	P
Neuseeland	Livest. Impro.					
Belgien	Univ. Liege					
Frankreich	INRA, Unceia	Holstein	9	1155	150	P, SCC
Normandie						
Montbéliarde						
Deutschland	ADR, 4Univ	Holstein	16			P, SCC
Finland	MTT	Ayrshire	12	493		P
Norwegen	As Univ	Norweg. Red	6	300	300	P
Schweden	Uppsala Univ	Swedish Red And White	13	515		P

P: Produktion, T: Typ, SCC: Somatic cell count, L: Langlebigkeit, F: Fertilität

[0008] Da die meisten Studien von kommerziellen Einrichtungen durchgeführt werden, werden zweifellos viele Ergebnisse nicht veröffentlicht. Grundsätzlich krankt jeder identifizierte QTL am gleichen Übel: Die meisten von ihnen sind nicht genauer als ca. 30–40 cM im Genom zu lokalisieren, prinzipiell ist mit den gegenwärtigen Suchstrategien nicht mehr als ca. 20 cM zu erreichen. Dabei gilt: je geringer die Heretabilität desto größer die Unsicherheit der Lokalisierung. So kommt es, daß trotz der gewaltigen Anstrengungen nur verhältnismäßig wenige QTL's als sicher molekular erfaßt gelten. Im Jahre 1998 konnten 14 QTL's, belegt durch meist mehrere einander bestätigende Publikationen, als molekularbiologisch abgesichert gelten.

[0009] Als Dilemma wird allerdings inzwischen angesehen, daß wegen der geringen Auflösung und Genauigkeit der jetzt verfügbaren Strategien zur QTL-Suche das erwünschte effektive MAS-Verfahren definitiv nicht in Sicht ist. Diese in zahlreichen Diskussionen des letzten Jahres auf höchstem wissenschaftlichem Niveau zum Ausdruck gebrachte Erkenntnis formuliert Michel Georges in seinem im August 1998 in Warschau gehaltenen Vortrag mit den Worten:

[0010] We are of the opinion that present mapping resolution is inadequate for the effective marker selection (MAS) and definitely insufficient for positional candidate cloning. Therefore, novel approaches have to be developed to refine the QTL map position.

[0011] Das ist vor allem deswegen eine schwierige Situation, weil gerade die zahlreichen Merkmale mit niedriger Heretabilität mit dem höchsten relativen Nutzen per MAS zu bearbeiten wären; wobei natürlich das zusätzliche experimentelle Problem die Tatsache ist, daß eine Kopplungsgruppe mit einem Merkmal niedriger Erblichkeit durch die vielfältigen und schwerwiegenden exogenen Einflüsse sehr viel schwerer aufzudecken ist als eines mit hoher Erblichkeit.

[0012] Als Perspektive ist also an dieser Stelle festzuhalten, daß eine Methode, die dieses prinzipielle Hindernis der zu ungenau lokalisierbaren Kopplungsgruppe überwindet, auch in einem ganz anderen qualitativen Sinne die molekulare, moderne Nutztierzucht verändern wird: Es werden dann nämlich auch die wirklich lohnenden QTL's bearbeitet werden können.

[0013] Eine Lösung des Problems ist zweifellos der Weg zum Kopplungsungleichgewicht. Findet man nämlich Marker, die im Kopplungsungleichgewicht zu den jeweiligen Verursachergenen stehen, so wird die Vorhersagesicherheit erheblich größer wegen der fehlenden Rekombination und man kann über die Familiengrenzen hinaus in den Bereich großer Populationen im oben geschilderten Sinne erfolgreich arbeiten (H. Bovenhuis et al. 1998).

[0014] Die gleiche Problematik findet sich im menschlichen Genom bei der Suche nach Verursachergenen für polygen vererbte Krankheiten. Da Defekte in mehreren, oft in sehr vielen Genen den Phänotyp der Krankheit erzeugen, sind diese so häufig, daß man sie Volkskrankheiten nennt. Obwohl mit Milliardenaufwand seit einigen Jahren die Kopplungsstrategie auf die Probleme von Diabetes, Herz-Kreislaufkrankungen oder erblicher Fettleibigkeit angewandt wird, gibt es noch bei keiner dieser Krankheiten einen umfassenden Erfolg. N. Risch et al. 1996 haben diesen Umstand aus der Sicht des Biostatistikers untersucht und ihre Folgerungen daraus gezogen: Um den gewünschten Effekt zu bekommen, müßte man gewaltige Zahlen von Patienten und Teile ihrer Familien mit dichten Markerkarten untersuchen. Die Zahlen sind so groß, daß man auch bei noch so großer finanzieller Anstrengung das Ziel nicht erreichen wird. Die Bemühungen sind jenseits der finanziellen Betrachtung auch deswegen zum Scheitern verurteilt, weil das notwendige, extrem sorgfältige Katalogisieren der phänotypischen Heterogenität bei den großen Probandenzahlen kaum zu bewältigen ist. Die Autoren ziehen in ihrer Situationsbeschreibung, die mit dem Titel "Future of Genetic Studies of Complex Human Diseases" überschrieben ist, den Schluß: Entweder die Technik der Genotypisierung muß einen Qualitätssprung machen oder man muß

eine Strategie finden, die direkt die Verursachergene oder die entsprechenden Kopplungsungleichgewichte aufzufinden gestattet.

[0015] Im Prinzip stellt die kürzlich mehrfach beschriebene Methode des Genomic-Mismatch-Scanning (GMS) eine solche Strategie dar. In dieser Schrift soll eine Gruppe von Verfahren zum Patent angemeldet werden, die auf einer ähnlichen Basis wie dieses GMS-Konzept beruhen. Diese Verfahren werden zur Lösung der oben geschilderten Probleme in der Tierzucht beitragen. Mit Seiner Hilfe kann man wie mit der Kopplungsstrategie Regionen in einem Genom lokalisieren, die für bestimmte genau umschriebene genetische Merkmale verantwortliche Gene enthalten. Die Suchoperation muß dabei in einer Population stattfinden, bei der das interessierende Merkmal, z. B. eine genetische Erkrankung, auf einen einzigen Vorfahren zurückzuführen ist. Das entsprechende Verursachergen und seine Umgebung im Genom bei allen Betroffenen der betrachteten Population stammen damit von diesem einen Vorfahren und sind in allen Sequenzeinheiten miteinander identisch. Im Angelsächsischen hat man dafür den Fachausdruck "Identity by Descent" (IBD) geprägt.

[0016] Mit der GMS-Methode hat man Kopplungsgruppen aber auch Kopplungsungleichgewichte isoliert. Dafür hat man zwei unterschiedliche Wege beschritten: Man definiert einen interessanten Phänotyp und studiert diesen mit Hilfe von Paaren, derer beide Individuen aus einer Familie stammen und einen eindeutig beschriebenen Verwandtschaftsgrad haben, wobei die einzelnen Paare aus jeweils genetisch heterogenen Populationen oder Familien stammen können.

[0017] Das Beschreiten des zweiten Weges setzt die Annahme voraus, daß in einer großen Population die interessanten phänotypischen Eigenschaften auf dieselben Vorfahren zurückgehen. Hier bildet man Paarungen, die aus nicht miteinander verwandten Individuen bestehen.

[0018] Auf dem ersten Wege isoliert man Kopplungsgruppen, auf dem zweiten Wege genomische Regionen, die ein Kopplungsungleichgewicht zwischen dem interessierenden Gen und einem aus mehreren polymorphen Genorten bestehenden Mikrohaplotyp darstellen.

[0019] Bisher hat man GMS-Protokolle nur auf monogen vererbte Erkrankungen angewendet. Davon gibt es in gewisser Weise eine Ausnahme: Um die Ausbeute und Zuverlässigkeit der Isolation von IBD-Fragmenten zu studieren, wurde eine Reihe von Mikrosatellitenloci in einem Experiment der GMS-Prozedur unterworfen. Damit hat man im Prinzip schon das Arbeiten an polygen vererbten Merkmalen simuliert.

[0020] Alle bisher publizierten Arbeiten enden mit der Lokalisierung bzw. der Bestätigung einer schon mit anderen Methoden vorgenommenen Lokalisierung der isolierten IBD's, also der GMS-Produkte.

[0021] Die Basis für alle diese GMS-Arbeiten stellt die Publikation von Nelson aus dem Jahre 1993 dar. Hier wurde zum ersten Male von einer erfolgreichen Anwendung dieses Protokolls auf die Lokalisierung von Genen bei der Hefe berichtet.

[0022] Die erste erfolgreiche Anwendung beim Menschen wurde 1997 veröffentlicht: F. Mirzayans identifizierte die chromosomale Region, die den Genort für die Iridogoniodysgenese (IGDA) beherbergen sollte. Der Befund wurde parallel mit der Kopplungsanalyse und der GMS erarbeitet. Bei der Krankheit handelt es sich um eine sehr seltene autosomal dominante Augenkrankheit. Für die Kopplungsanalyse untersuchte man eine Familie mit 80 Mitgliedern, von denen 40 erkrankt sind. Das ganze Genom ist dabei mit 300 Mikrosatelliten untersucht worden (Mirzayans et al 1995). Danach war das IGDA-Gen auf 6p lokalisiert. In einer nächsten Lokalisierungsrunde verwendete man 29 Mikrosatelliten-Marker des Chromosom 6. Der Genort wurde damit auf 6p25 weiter eingegrenzt.

[0023] Mit einem im Wesentlichen gegenüber dem von Nelson unveränderten GMS-Protokoll isolierte man IBD-Fragmente von zwei Cousins S. Grades aus der erwähnten Familie. PCR mit den entsprechenden Primern zeigte, daß von den 7 erhaltenen positiven PCR-Signalen (7 Mikrosatelliten) 5 aus einer chromosomalen Region stammten, die auch mit der konventionellen Kopplungsanalyse eine signifikante Kopplung aufwiesen. Die hintereinander liegenden IBD-Fragmente mit positivem PCR-Signal und signifikanter Kopplung erstreckten sich über einen Bereich von 6,9 cM.

[0024] Als interne Kontrolle des GMS-Experimentes wurde parallel zu Chromosom 6 das Chromosom 12 untersucht. Letzteres ergab kein positives Signal.

[0025] Das entscheidende Motiv für die Arbeit von Mirzayans et al war der Versuch, GMS als einen Ansatz zu etablieren, mit dessen Hilfe der Genort einer seltenen Erkrankung ohne Durchführung und Auswertung tausender von Mikrosatelliten-Typisierungen zu lokalisieren ist. Dieser Versuch ist zweifellos gelungen. Entscheidende Voraussetzung für solch ein erfolgreiches Experiment an einem einzelnen Paar von Indexpersonen ist die Auswahl des Verwandtschaftsgrades der beiden Individuen. Qualitativ ist dieser Zusammenhang klar: Sind die beiden Individuen zu nah verwandt, muß man mit sehr großen IBD's rechnen, im Grenzfall (s. weiter unten) 80% eines Chromosoms oder mehr, daneben wird es IBD's geben, die nichts mit dem interessierenden Merkmal zu tun haben. Ist dagegen der Verwandtschaftsgrad zu gering, kann der IBD-Bereich wegen seiner geringen Größe mit dem Lokalisierungsinstrument nicht mehr entdeckt werden, weil dieser eine zu geringe Auflösung hat (Der von Mirzayans et al verwendete Markersatz hat eine Auflösung von 10 cM). Daraus erhellt die enorme Bedeutung des Lokalisierungsinstruments im GMS-Protokoll.

[0026] Offensichtlich gibt es bei der Isolierung von IBD's falsch-negative und falsch-positive Resultate. Die falsch negativen werden nicht diskutiert; es ist aber nicht überraschend, daß es solche gibt. Nach dem ursprünglichen Protokoll setzt man 5 µg DNA vor jedem Indexindividuum ein, Mirzayans setzen immerhin 100 µg ein, trotzdem ist auch bei letzterem Experiment die Ausbeute vermutlich nicht höher als 100 ng. Dabei handelt es sich aber um eine Population von ca 500-1000 verschiedenen Fragmenten, so daß man nicht sicher sein kann, daß alle amplifiziert werden können. Darüber hinaus kann das Reannealing sequenzabhängig unvollkommen sein und so ein mutationsbedingtes Mismatch vortäuschen. Für falschnegative Ergebnisse gibt es einige Erklärungen – so weiß man, daß Mut S keine C-C-Mismatches erkennt, und man weiß, daß Mut S zur Erkennung eines Mismatches ein methyliertes GATC-Motiv in dessen Nachbarschaft benötigt. Letzteres ist natürlich nicht immer gegeben. Ebenso ist die Funktion von Mut L, H nicht ganz klar.

[0027] Als Ergebnis dieses Komplexes theoretischer und experimenteller Arbeiten, die letztlich in die Arbeit von Mirzayans einmünden, ist festzuhalten, daß ein monogenes, erbliches Merkmal in einem einzigen GMS-Experiment im menschlichen Genom zu lokalisieren ist. Die Ausdehnung der nachgewiesenen Kopplungsgruppe ist dabei 6,9 cM. Die Markerkarte hat eine Auflösung von 10 cM und es stand eine Index-Familie zur Verfügung, die die Analyse von Cousins S. Grades zuließ.

[0028] Das Verfahren zeigt in erheblichem Maße falschpositive und falschnegative Signale, die höchstwahrscheinlich mit der im zur Verfügung stehenden GMS-Protokoll vorgegebenen Enzymausstattung und den vorgeschlagenen Anreicherungs-schritten unvermeidlich sind. Deswegen erscheint es aussichtslos, die bislang beschriebenen Arbeitsvorschriften auf multifaktorielle bzw. auf polygen vererbte Merkmale anzuwenden.

[0029] Die zweite mit der GMS verfolgte genetische Zielrichtung ist der Versuch direkt Bereiche zu isolieren, die Kopplungsungleichgewicht zeigen. Ein Beispiel für diesen Ansatz ist die Arbeit von Vivian G. Cheung (1998a). Hier wird das die Krankheit verursachende Gen des autosomal-rezessiv vererbten congenitalen Hyperinsulinismus lokalisiert. Die Krankheit kommt mit einer vergleichsweise hohen Frequenz bei den Ashkenazi Juden vor. Deswegen kann man in dieser Bevölkerungsgruppe von einem Foundereffekt ausgehen. Das Gen kodiert für den Sulfonylharnstoffrezeptor. Es war mit Hilfe konventioneller Methoden auf dem Chromosom 11p15.1 lokalisiert worden.

[0030] Es wurden insgesamt 8 betroffene Personen aus 4 Familien in zwei unterschiedlichen Ansätzen untersucht. Im ersten wurde aus Geschwistern einer Familie Paare gebildet, im zweiten geschah die Paarbildung zwischen Betroffenen aus unterschiedlichen Familien. Die Lokalisierung der IBD's wurde auf einem Glaschip vorgenommen, der nach chromosomaler Lokalisierung geordnete inter-Alu-Banden von YAC- und PAC Banken von Chromosom 11 enthielt.

[0031] Bei den Geschwistern erhielt man so eine IBD-Strecke, die ca 80% des Chromosom 11 enthielt (theoretische Erwartung ist 75%). Die Auflösung der Bank betrug außerhalb des ja vorher bekannten Genortes des SUR1-Gens 2 cM., in der Gegend von 11p15.1 war sie ca 0,25 cM. Es werden die Hybridisierungsdaten der aus 9 unterschiedlichen Paarungen unverwandter Personen gezeigt; Die Signale erstrecken sich über rund 2,5 cM mit unterschiedlichen Ausdehnungen bei den einzelnen Paarungen, dagegen haben 8 Paarungen ein Signal bei einem Klon, in den beiden benachbarten Klonen finden sich jeweils Signale von 7 Paarungen. Daraus leitet man im günstigsten statistischen Falle ein Lokalisierungsge-nauigkeit von 500 bp ab. Bei 4 Paarungen findet man einige positive Signale völlig außerhalb der SUR1-Region.

[0032] Das wichtigste Ergebnis im Zusammenhang mit dem vorliegenden Projekt ist sicher die Erkenntnis, daß das Kartieren von Genorten über Kopplungsungleichgewicht mit GMS möglich ist. V. Cheung et al. räumen ein, daß ohne die vorher bekannte Lokalisation des Gens einige Paarungen mehr hätten untersucht werden müssen, um statistisch völ-lig abgesicherte Ergebnisse zu erhalten.

[0033] Das Ermitteln von Kopplungsungleichgewicht ist überaus bedeutsam: Die bisher mit Hilfe der Kopplungsstra-tegie identifizierten QTL's sind häufig nicht brauchbar, weil die mit den interessierenden Merkmalen gekoppelten Mar-ker-Loci eben nur Kopplung und keine Assoziation zeigen, das bedeutet, man kann sie nur für eine indirekte Genanalyse benutzen. Das heißt natürlich auch, daß sie für einen umfassenden Gebrauch nutzlos sind. Bis man aber zu assoziierten Haplotypen vorgedrungen ist, wird, wie man von den entsprechenden Problemen aus der Humangenetik weiß, sehr Lange dauern. Es besteht der Verdacht, daß es in vielen Fällen unmöglich ist, wenn man bei der jetzigen Strategie bleibt (Risch et al 1996).

[0034] In V. Cheung's Arbeit wird in besonderem Maße klar, welche entscheidende Rolle die schnelle und genaue Lo-kalisierung spielt und daß dafür augenblicklich der geeignetste Ansatz die Chip-Hybridisierung einer guten genomischen Bank mit hoher Auflösung ist. Auch hier lassen allerdings die beschriebenen Verfahren ein wesentliches Problem unge-löst, nämlich das der verfügbaren DNA-Menge: Alle theoretischen und praktischen Betrachtungen der Ausbeute am Ende einer GMS-Prozedur erlauben keinen allzu großen Optimismus; man wird sich immer an der Grenze der Nachweis-barkeit der gegenwärtig zur Verfügung stehenden Amplifikationsstrategien bewegen, wie weiter unten mit einigen Zah-len belegt wird.

[0035] Es gibt in der internationalen wissenschaftlichen Gemeinschaft keinen Zweifel daran, daß eine in allen Punkten funktionierende GMS oder ähnliche Genom-Scanning-Methode bei der Aufklärung komplexer genetischer Merkmale die jetzige Kopplungsstrategie ablösen wird. In ihrer modellhaft angelegten Studie prüften V. Cheung (1998b) eine Reihe von Parametern, die den Erfolg der Anwendung einer GMS-Prozedur auf polygene Merkmale charakterisieren könnten.

[0036] In der Arbeit sind insgesamt 25 getrennte GMS-Experimente an 14 unterschiedlichen Großeltern-Enkelpaaren aus CEPH-Familien durchgeführt worden. Dabei sind die einzelnen Paare mit bis zu 33 verschiedenen Mikrosatelliten-Loci untersucht worden. Die Ergebnisse sind in Wiederholungsexperimenten reproduziert worden. Der hauptsächliche Parameter ist der Anreicherungs-faktor des IBD-Allels. Jeder Genort weist ein Allel auf, welches nur entweder von dem väterlichen Großvater oder von der väterlichen Großmutter ererbt werden sein kann. Dieses Allel ist das jeweilige IBD-Fragment oder IBD-Allel. An zwei Positionen des GMS-Experimentes wird eine quantitative Messung (Fläche unter dem Fluoreszenzsignal-Peak des Mikrosatellitenlocus in der Messung mit einem automatischen Sequenziergerät) durch-geführt; zum einen an der Gesamtpopulation der Heterohybride nach der Restriktions- und Exo III-Verdauung der Ho-mohybride und und zum anderen an den verbliebenen Heterohybriden nach Abtrennung der die Fehlpaarungen-enthal-tenden Heterohybridfraktion. Man definiert dann den gesuchten Anreicherungs-faktor wie folgt: Den Nenner dieses Ver-hältniswertes bildet das Verhältnis aus IBD-Allel zum Nicht-IBD-Allel in der Gesamtpopulation der Heterohybride, den Zähler bildet das entsprechende Verhältnis in der komplett matchenden GMS-Fraktion. Man unterscheidet drei Gruppen von Anreicherungsvermögen, nämlich > 10 ; $2-10$; < 2 . Dies über alle verwendeten Mikrosatelliten und alle Genotypi-sierungen (122) betrachtet, ergab: 29% hohe Anreicherung, 45% mittlere Anreicherung und 26% niedrige Anreicherung. Bezieht man die Beobachtung dagegen auf die einzelnen verwendeten Genorte (33), so erkennt man unveränderlich in je-dem Experiment bei 7 Markern (21%) hohe Anreicherung, bei 4 Markern (12%) mittlere Anreicherung und bei 7 Mar-kern (21%) niedrige Anreicherung. Bei 9 Markern erhielt man im Laufe der wiederholten Experimente variable Anrei-cherung, d. h. niedrige bis hohe. Bei 6 Markern erhielt man gar keine Anreicherung. Offensichtlich ist die Gesamtproze-dur in hohem Maße reproduzierbar. Die Anreicherung, also der Erfolg der GMS-Prozedur ist sequenz- und damit genort-abhängig, da rund 18% der verwendeten Markerloci keine Anreicherung zeigen. Daraus ist allerdings noch keine allge-meingültige Effizienzbewertung des Verfahrens abzuleiten, da ausschließlich Pst I-Fragmente untersucht wurden. Die Auswahl der Restriktionsspaltung stellt ein Optimierungsproblem dar, welches sich mit mindestens zwei Einflußgrößen auseinanderzusetzen hat: Jedes DNA-Molekül, das im Laufe der GMS entfernt werden soll, muß mindestens ein GATC-Motiv tragen – je kleiner die Fragmente sind, desto höher ist die Wahrscheinlichkeit, daß kein GATC vorhanden ist. Um-gekehrt finden in der FPERT-Hybridisierung sicherlich komplementäre DNA-Stränge umso schlechter zueinander je

mehr repetitive Fragmente sie enthalten. Die Wahrscheinlichkeit dafür steigt natürlich mit der Fragmentlänge. Fragmente > 20 kb werden offensichtlich schon während des Reannealings abgebaut.

[0037] Überhaupt ist mit den gegenwärtigen Protokollen die Ausbeute der verschiedenen Präparationsstufen im Vergleich zum ursprünglichen Hefeexperiment unbefriedigend. Dies ist nicht überraschend, weil das Hefegenom 240 mal weniger komplex als das menschliche ist, weil es viel weniger repetitive Sequenzen enthält und weil es weit mehr natürliche Sequenzpolymorphismen trägt als das menschliche Genom. Das alles führt nämlich dazu, daß die FPERT-Stufe weniger effizient ist: Während man bei einem GMS-Experiment an Hefe rund 50% der eingesetzten DNA nach der FPERT als Doppelstrangmoleküle wiederfindet (jeweils ca 50% Heterohybride und Homohybride), sind es beim Menschen nur 10%. Das bedeutet, in einem typischen Experiment gewinnt man ca 1 µg DNA. Nach Verdauung der Homohybride bleiben ca 500 ng. Daraus werden die IBD-Fragmente abgetrennt, dabei kann es sich natürlich nur noch um einige ng handeln – umso weniger je schärfer die GMS selektiert. Auch die vorgeschlagene und in einem Falle durchgeführte PCR-Amplifikation über inter-Alu-Motive befindet sich damit im problematischen Grenzbereich von 1 ng pro Fragment (Nelson et al 1989).

[0038] Das in dieser Patentanmeldung beschriebene Verfahren dient der Isolation von IBD-Fragmenten bei polygen vererbten Merkmalen. Das Verfahren weist entscheidende Unterschiede und Verbesserungen gegenüber allen vorher beschriebenen Protokollen auf.

Literatur

- 20 Botstein D, White R, Skolnik M, Davies RW, Am J. Hum. Genet, 32: 314–331, (1980)
Nelson SF, McCusker JH, Sander MA, Y. Kee, Modrich PO, Brown P Nature Genetics 4, 11–18, (1993)
Mirzayans F, Mears AJ, Guo W, Pearce WG, Walter MA Am J Hum Genet 61, 439–448, (1997)
Mirzayans F, Pearce WG, MacDonald IM, Walter MA Am J Hum Genet 57 (3), 539–48 (1995)
Risch N, Merikangas K Science 273, 1516–1517 (1996)
25 Nelson, U. L., Ledbetter, SA, Corbo L, Victoria NF, Ramirez-Solis R, Webster TD, Ledbetter DH, Caskey CT, Proc. Natl. Acad. Sci USA 86: 6686–6690 (1989)
Bovenhuis H, Spelman RJ, 49th Annual Meeting European Association for Animal Production, August 24–27 (1998), Warsaw, Poland
Georges M (1998) Mapping genes underlying production traits in livestock. P 77–101. In Animal Breeding: Technology for the 21st century. Ed. J. Clark. Harwood Academic Publishers.
30 Georges M, Nielsen D, Makinon M Genetics 139, 907–920, (1995)
Geldermann H, Theor Appl Genet 46, 319–330, (1975)
Weiler JL, Kashy Y, Soller M, J Dairy Sci, 73, 2525–2537, (1990)
Cheung V, Genomics 47, 1–6, (1998a)
35 Cheung V, Gregg JP, Gogolin-Ewns KJ, Bandong J, Stanley CA, Baker L, Higgins MJ, Nowak NJ, Shows TB, Ewns WJ, Nelson SF, Spielman RS Nature Genetics 18, 3, 225–230, (1998b)

Problemstellung

40 [0039] Die bis heute verwendeten Verfahren (GMS) zur Isolierung von IBD (Identity by Descent Bereichen des Erbmaterials weisen eine Reihe von prinzipiellen Schwierigkeiten auf, die insbesondere die Erfolgsaussichten ihrer Anwendung auf komplexe genetische Merkmale zweifelhaft erscheinen lassen.

- Die zur physikalischen Abtrennung der Homohybride erforderliche Methylierung der DNA des einen Hybridisierungspartners ist nur schwer steuerbar, weil wesentliche Teile jedes Genoms methyliert sind. Das kann zu erheblichen Ausbeuteverlusten beim Restriktionsverdau der nichtmethylierten DNA führen.
- Es ist nicht klar, inwieweit der Restriktionsverdau der methylierten Fraktion vollständig ist.
- Der Mut S, H, L-Komplex fordert für seine Reaktion der Erkennung und des Schnitts eine Methylierung in Nachbarschaft eines GATC-Sequenzmotivs.
- 50 – Mut S erkennt nicht alle möglichen Mismatch-Kombinationen.
- Der Enzymkomplex fordert die Nachbarschaft eines GATC-Motivs.
- An zwei Stellen des beschriebenen GMS-Verfahrens muß einzelsträngige DNA mit BNCD isoliert werden. Dies sind verlustreiche reinigungsschritte.
- Die größte Beschränkung stellt die geringe Ausbeute der gesamten Prozedur dar.

Beschreibung des zu schützender Verfahrens

1. Die DNA von zwei Indexindividuen wird extrahiert und gereinigt.
2. Beide DNA-Fraktionen werden mit dem gleichen Restriktionsenzym verdaut – im weiter unten (Abbildung) beispielhaft erörterten Fall handelt es sich um das Restriktionsenzym Eco RI.
- 60 3. Die restringierte DNA von Individuum 1 wird mit einem Linker versehen, der an beiden Seiten Fragmentenden ohne Überhang erzeugt (s. Abbildung).
4. Die restringierte DNA von Individuum 2 wird ebenfalls mit einem Linker versehen, der an beiden Seiten Fragmentenden ohne Überhang erzeugt (s. Abbildung).

65 [0040] Für beide Individuen sind die Linker so gestaltet, daß für Heterohybride aus der DNA beider Individuen keine Selbstligation zur Ringbildung möglich ist, daß sie aber Ringbildung mit entsprechend geschnittenem Vektor zulassen.

5. Die DNA beider Individuen wird denaturiert.
6. Die denaturierte DNA beider Individuen wird gemischt.
7. Die Mischung wird der sogenannten FPERT-Reaktion zur Rückbildung doppelsträngiger DNA ausgesetzt.
8. Am Ende dieser Reaktion besteht die Reaktionsmischung aus drei Molekülpopulationen, nämlich aus den zurückgebildeten doppelsträngigen DNA-Molekülen von Individuum 1 (Homohybride), den doppelsträngigen DNA-Molekülen von Individuum 2 (Homohybride) und den doppelsträngigen DNA-Molekülen, die aus einem von Individuum 1 herrührenden DNA-Strang und einem von Individuum 2 herrührenden DNA-Strang bestehen (Heterohybride).
9. Ein geeigneter Plasmidvektor wird einem Doppelverdau z. B. mit den Restriktionsenzymen Nco I und dem Restriktionsenzym Nsp I ausgesetzt. Dabei entstehen auf dem gleichen Strang in entgegengesetzter Orientierung überstehende GTAC-Enden (s. Abbildung).
10. Die Mischung der neu entstandenen Doppelstrangmoleküle wird mit der geschnittenen Vektor-DNA zusammengegeben.
11. Nur die Heterohybride können mit den Vektormolekülen einen Ringschluß bilden. Die DNA-Fragmente werden mit einer Ligase-Reaktion kovalent verknüpft.

[0041] Nach dieser Reaktion sind also drei Molekülpopulationen entstanden – die Homohybride der beiden Individuen 1 und 2 und die Heterohybriden, die aus einem Strang von Individuum 1 und dem komplementären Strang von Individuum 2 bestehen.

[0042] Die Heterohybride ihrerseits bestehen aus der sehr großen Gruppe der Moleküle, die einzelne Fehlpaarungen aufweisen und der wesentlich kleineren Molekülgruppe, die keinerlei Fehlpaarungen haben. Letztere Gruppe enthält die interessanten Fragmente, nämlich die IBD-Bereiche.

12. Mit Hilfe einer CEL I-Reaktion werden alle Basenfehlpaarungen erkannt und an einem Strang aufgeschnitten.
13. Ausgehend von den im Schritt 13 geöffneten, ungepaarten Bereichen der Fehlpaarungen werden dann mit Hilfe eines Exonuklease III Schrittes diese Moleküle abgebaut.
14. Ohne weitere Vorreinigung wird der gesamte Ansatz zur Transformation geeigneter Bakterien vorbereitet.
15. Die transformierten Bakterien werden in entsprechendem Selektionsmedium ausplattiert.
16. Die DNA der einzelnen Plasmidpräparationen stellt jeweils ein Fragment dar, welches "identical by descent" (IBD) ist.
17. Die erhaltene DNA wird auf einem DNA-Chip lokalisiert.

Anwendungsbeispiel

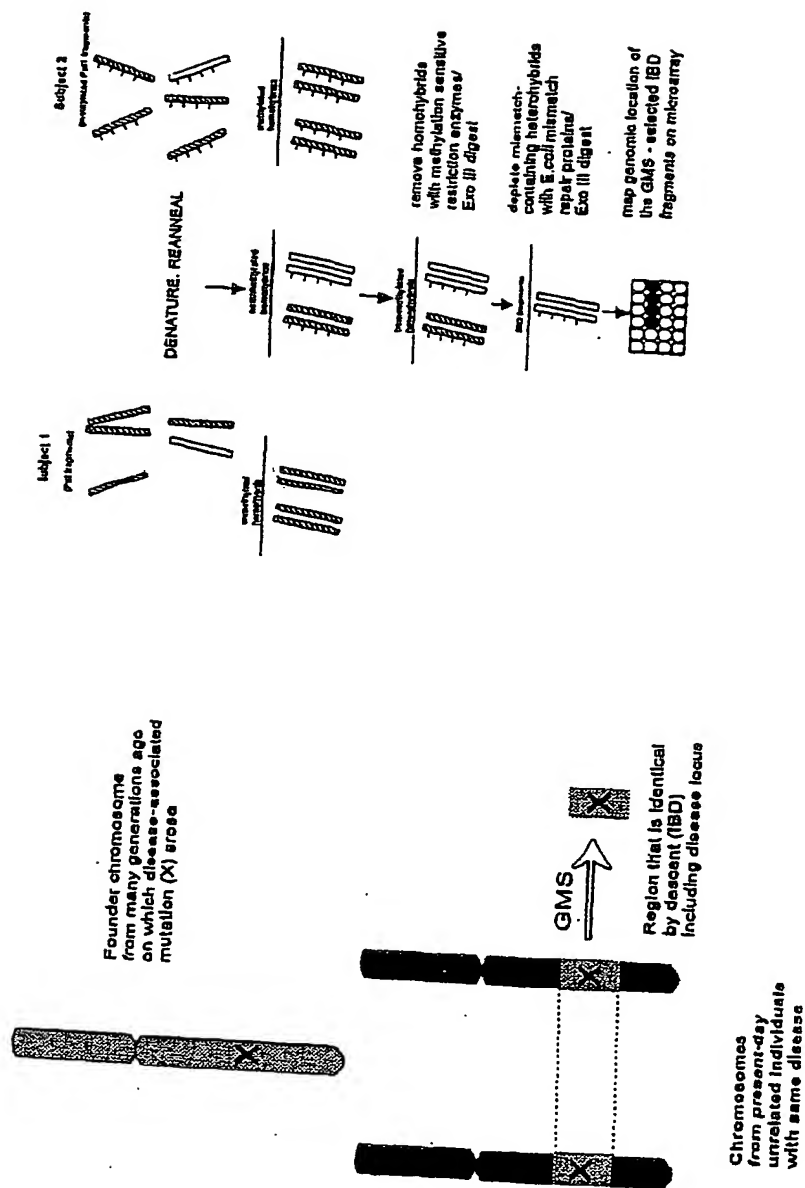
[0043] Eine verhältnismäßig einfache Anwendung ist die Identifizierung und die Isolierung des Gens für die ererbte Afterlosigkeit beim Hausschwein. Mit der von uns beschriebenen Prozedur erhält man mehrere IBD-Fragmente. Diese Fragmente lassen sich mit dem DNA-Chip lokalisieren. Mit hoher Wahrscheinlichkeit ergeben sich dabei mehrere Genorte. Diese werden auf ihre Aussagefähigkeit hin geprüft mit betroffenen und nicht betroffenen Tieren. Die informativen Genorte können für diagnostische und weitergehende wissenschaftliche Zwecke genutzt werden.

Patentansprüche

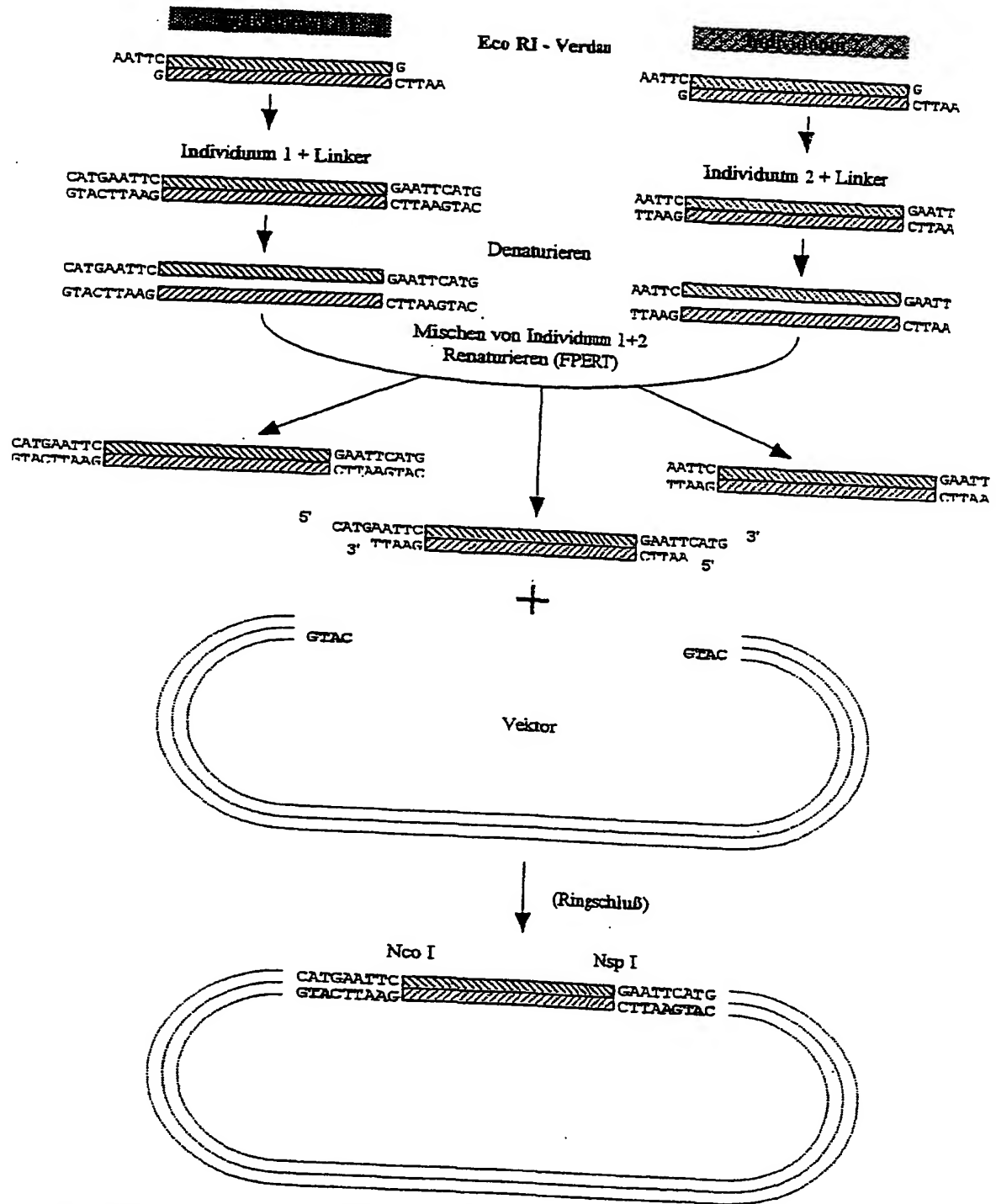
1. Wir melden hier Schutzrechte für ein Verfahren an, mit dessen Hilfe bei nicht miteinander verwandten sowie bei miteinander verwandten Individuen Bereiche des Genoms isoliert werden können, die Kandidatengenbereiche enthalten, welche sich im Kopplungsungleichgewicht mit ihrer engeren DNA-Umgebung befinden. Die Methode ist **dadurch gekennzeichnet**, daß sie die Isolation von Kandidatengen gestattet, die komplexe genetische, also polygen vererbte Merkmale steuern.
2. Die Methode ist dadurch gekennzeichnet, daß sie die mit Restriktionsenzymen (im hier erläuterten Fall ist EcoR I beispielhaft angegeben) geschnittenen DNA-Proben von zwei Indexindividuen mit jeweils unterschiedlichen Linkern versieht, die die folgenden Eigenschaften haben:
 - 2.1 Sie erzeugen an beiden Enden der Restriktionsfragmente Enden ohne Überhang.
 - 2.2 Wenn sich als Ergebnis der im nächsten Schritt ablaufenden Reaktion Heterohybride aus den einander komplementären, jeweils zu Individuum 1 oder 2 gehörenden DNA-Strängen bilden, so ergeben sich an beiden Fragmentenden Überhänge mit folgenden Eigenschaften:
 - 2.2.1 Sie befinden sich am gleichen Strang
 - 2.2.2 Ihre Sequenz ist am 5'-Ende 5'CATG3' und am 3'-Ende 5'CATG3'
 Wird ein anderes Enzym als Eco RI benutzt, so haben die Überhänge eine andere Sequenz; diese weist aber die gleiche Orientierung auf.
3. Die Methode ist dadurch gekennzeichnet, daß sie die Denaturierung dieser Fragmente beinhaltet
4. Die Methode ist dadurch gekennzeichnet, daß sie die Mischung der so erhaltenen einzelsträngigen DNA-Fragmente von beiden Individuen beinhaltet
5. Die Methode ist dadurch gekennzeichnet, daß im nächsten Schritt Pufferbedingungen eingestellt werden, die die Renaturierung von doppelsträngigen DNA-Molekülen zuläßt (FPERT-Reaktion).
6. Die Methode ist dadurch gekennzeichnet, daß drei Molekülpopulationen entstanden sind, von denen die erste die wieder hergestellten Doppelstrangmoleküle (Homohybride) von Individuum 1 darstellen, die zweite die wiederhergestellten Doppelstrangmoleküle (Homohybride) von Individuum 2 und die dritte die Population der zueinander komplementären, jeweils dem einen und dem anderen Individuum zugehörenden Doppelstrangmoleküle (Heterohybride) darstellen.

7. Die Methode ist dadurch gekennzeichnet, daß Pufferbedingungen eingestellt werden, die einen Ringschluß der entstandenen Moleküle zulassen
8. Die Methode ist dadurch gekennzeichnet, daß für das hier gewählte Beispiel der Eco RI-Verdauung ein geeigneter Klonierungsvektor einem Doppelverdau mit den Restriktionsenzymen Nco I und Nsp I ausgesetzt wird.
9. Die Methode ist dadurch gekennzeichnet, daß die so behandelte Vektor-DNA gemischt wird mit der Reaktionsmischung aus Schritt
10. Die Methode ist dadurch gekennzeichnet, daß in Folge der spezifischen Konstruktion der Linker nur die Heterohybride den Ringschluß mit den Vektormolekülen vollziehen.
11. Die Methode ist dadurch gekennzeichnet, daß mit Hilfe einer Ligase-Reaktion die ringförmig miteinander verknüpften DNA-Fragmente kovalent verbunden werden.
12. Die Methode ist dadurch gekennzeichnet, daß die ringförmigen Moleküle aus zwei Populationen bestehen, nämlich aus derjenigen mit ein oder mehr Basenfehlpaarungen und aus derjenigen ohne jede Fehlpaarung. – Letztere Gruppe enthält die IBD (identity by descent) – Bereiche.
13. Die Methode ist dadurch gekennzeichnet, daß diese Reaktionsmischung unter geeigneten Pufferbedingungen mit dem Enzym CEL I versetzt wird.
14. Die Methode ist dadurch gekennzeichnet, daß CEL I jede Fehlpaarung erkennt und einen der beiden DNA-Stränge an der Position der Fehlpaarung schneidet.
15. Die Methode ist dadurch gekennzeichnet, daß dieser Reaktionsmischung nach Einstellen geeigneter Pufferbedingungen die Nuklease Exo III hinzugefügt wird. – Als Ergebnis werden die durch den CEL I-Verdau partiell einzelsträngigen Ringe abgebaut.
16. Die Methode ist dadurch gekennzeichnet, daß diese Reaktionsmischung zur Transformation geeigneter Bakterien eingesetzt wird.
17. Die Methode ist dadurch gekennzeichnet, daß die transformierten Bakterien nach Selektion und Vereinzelung kultiviert werden, und daß die per Plasmidpräparation gewonnenen DNA-Fragmente – sie enthalten die IBD-Bereiche – isoliert und im Genom lokalisiert werden.

Hierzu 2 Seite(n) Zeichnungen



Die obige Abbildung beschreibt schematisch den Gang des Genomic-Mismatch-Scanning-Verfahrens.



Schematische Darstellung des zu schützenden Verfahrens